

## LETTERS

### The reliability of cancer registry records

In their recent paper, which compared information obtained from the cancer registry with data abstracted from medical records, Pollock and Vickers concluded that "disagreements over date of diagnosis will bias survival data, disagreements over site will affect incidence data and trends, and disagreements over treatment undermine the case for using registry data to evaluate care."<sup>1</sup> To what extent are these conclusions supported by data included in their paper?

For cases with data available both from the cancer registry and from case notes, dates of diagnosis agreed to within 30 days or less for 305 of 385 cases (79%). For date of death, there was exact agreement between the two data sets for 245 of 270 cases (91%). These findings show some lack of precision in recording date of diagnosis, particularly for cases in which registration was initiated from the death certificate. This paper has not presented data which show systematic bias from misclassification of date of diagnosis or date of death, either for the sample as a whole or for subgroups. In our own study of bladder cancer<sup>2</sup> we found that there was exact agreement between data in the cancer registry and case notes for date of first operation in 83% of cases. For discordant cases the median difference was -1 day (interquartile range -10 to 3 days). Similarly for date of death there was exact agreement in 93% of cases and for discordant cases the median difference was 1 (-3 to 5) days. These data showed a high level of agreement between case notes and cancer registry data. In a small proportion of cases the date of first operation or date of death were recorded imprecisely but there was no evidence of a systematic bias that might influence survival estimates.

Turning to data presented for tumour site, the rate of discordance is similar to that reported for other registries.<sup>3</sup> After excluding cases which the reviewers were unable to classify topographically (which are difficult to interpret as the authors do not give the codes used to sample from the cancer registry), data given in their table 4 are consistent with a kappa statistic of 0.86 (95% confidence interval 0.76 to 0.96). According to criteria conventionally used to evaluate the quality of epidemiological data, this is consistent with excellent agreement between the two data sources.<sup>4</sup> No data were presented to show that the quality of recording of details changed over time in a manner that might influence the assessment of secular trends; in fact, the rate of discordance was the same in both years studied.

Although it is clear that cancer registries are not currently in a position to document cancer treatment in a comprehensive way, this does not undermine the

case for using cancer registrations as a sampling frame for evaluative research. Cancer registry data have been shown repeatedly to be of value in comparing the outcome of cancer at international,<sup>5</sup> national,<sup>6</sup> and local levels.<sup>7,8</sup> Recognition of a central role for population based disease registers in public health research has led to the increasing use of registers in the study of other conditions.

Several other points require qualification. The authors seem to regard their own data as the reference material but they have not evaluated the reliability of their own data abstraction and it is possible that an experienced cancer registry clerk may perform better than a less experienced researcher.<sup>9</sup> Because the authors only retrieved case notes for 62% of cases provided by the cancer registry, reliability was studied for 416 cases and not the 673 cases mentioned in the title of the paper. The term "death certificate only registration" is used inconsistently and the term is used in the abstract to include cases for which details of diagnosis and treatment were available from the registry.

Cancer registrations, in common with other sources of routinely collected information, are not free from error and every effort should be made to allow cancer registries to improve the quality of their work. To observe that errors exist is unremarkable, it is more important to estimate the impact of such errors on subsequent data analyses. Evaluation of cancer registry data should be performed with care. The conclusions of this paper receive only limited support from the data presented.

MARTIN GULLIFORD

Department of Public Health Medicine,  
UMDS, Guy's and St Thomas's  
Medical and Dental Schools,  
St Thomas's Hospital,  
London SE1 7EH

- 1 Pollock AM, Vickers N. Reliability of data of the Thames cancer registry on 673 cases of colorectal cancer: effect of the registration process. *Quality in Health Care* 1995;4: 184-9.
- 2 Gulliford MC, Bell J, Bourne HM, Petrukevitch A. The reliability of cancer registry records. *Br J Cancer* 1993;67: 819-21.
- 3 Brewster D, Crichton J, Muir C. How accurate are Scottish cancer registration data? *Br J Cancer* 1994;70:954-9.
- 4 Fleiss JL. *Statistical methods for rates and proportions*, 2nd ed. New York: John Wiley, 1981:212-36.
- 5 Dillner L. Britain's breast cancer services need overhaul. *BMJ* 1995;311:211-2.
- 6 Silman AJ, Evans SJW. Regional differences in survival from cancer. *Community Medicine* 1981;3:291-7.
- 7 Chouillet AM, Bell CJM, Hiscox JG. Management of breast cancer in south east England. *BMJ* 1994;308:168-71.
- 8 Gulliford MC, Petrukevitch A, Burney P. Survival with bladder cancer: evaluation of delay in treatment, type of surgeon, and modality of treatment. *BMJ* 1991;303: 437-40.
- 9 Gulliford MC, Bell J, Bourne H, Petrukevitch A. The reliability of cancer registry records [letter]. *Br J Cancer* 1993;68:1046.

**AUTHORS' REPLY** - We thank Gulliford for his comments. We shall answer each in turn.

He is right to say that our paper presents no evidence of systematic bias in the

recording of dates of diagnosis and dates of death. Our objectives (as stated in the abstract and introduction to our paper) were "to measure the reliability of data collected by the Thames cancer registry and to identify factors in the registration process affecting reliability."<sup>1</sup>

The large differences between his study and ours on absolute rates of agreement for date of diagnosis and date of treatment are in part due to our inclusion of "death certificate only" registrations (DCOs).<sup>2</sup> We consider it essential to have included these. Gulliford and colleagues, confining themselves to 1982 data, found that DCOs accounted for only 12 (2%) cases and they excluded them from their sample. Historically, DCO registrations were based only on those cases that remained unregistered after intensive and extended searches for information on tumour site and date of diagnosis. Since 1983, a rapid increase has taken place in Thames DCO rates. In a study of DCO registrations in the Thames Regions registered between 1987 and 1989, we found that 24% of all malignant neoplasms were registered by DCO.<sup>3</sup> This compares with a national figure of less than 4%. The registry has explained this rise to be a result of the decision taken in 1983 (for financial reasons) not to follow up cases who died at home.<sup>4,5</sup> We thought it unlikely that all of the 150 (22%) DCO cases in our sample had had no contact with hospital services and for this reason we requested notes on DCO cases. We retrieved notes on 66 (44%) cases recorded by the registry as DCO cases, 12% of which had a date of diagnosis preceding the date of death by more than a year.

We found 49 disagreements over ICD code (12% of our sample). This figure fell to 33 (8% of our sample) once DCO registrations were excluded. Although we were unable to consult the paper by Brewster *et al* cited by Gulliford (it was published in the same month that our paper was accepted for publication), we did cite other papers which report similar findings: West found error and omission rates of 7.5% for ICD code and Lapham and Waugh reported disagreements of 4% when measuring the accuracy of Scottish cancer registry data against pathology reports.<sup>6,7</sup> To find that the proportion of disagreements reported is comparable with those of other studies is unremarkable. Our concern lies with the fact that that proportion may be high enough to obscure changes in incidence and survival over time. This seems to us more relevant than a kappa statistic showing excellent agreement. The rise in DCO cases over time may also have affected the reliability of tumour data; several studies have cast doubt on the accuracy of tumour site as recorded on death certificates.<sup>8</sup> Cases were requested from the Thames cancer registry with an ICD code of 153 or 154.

Like Gulliford we are strong supporters of the cancer registration system. However, we have two concerns about using the Thames registry as a sampling frame. Firstly, the exclusion of DCO cases from the sample can lead to bias. We have